Catalogue no. 82-003-X
ISSN 1209-1367

Health Reports

Assessment of the quality of the childhood physical abuse measure in the National Population Health Survey

by Margot Shields, Wendy Hovdestad and Lil Tonmyr

Release date: May 20, 2015

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at infostats@statcan.gc.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following toll-free numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-877-287-4369

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "About us" > "The agency" > "Providing services to Canadians."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard table symbols

The following symbols are used in Statistics Canada publications:

.	not available for any reference period
..	not available for a specific reference period
...	not applicable
0	true zero or a value rounded to zero
0^s	value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
^{p}	preliminary
^{r}	revised
x	suppressed to meet the confidentiality requirements of the *Statistics Act*
^{E}	use with caution
F	too unreliable to be published
*	significantly different from reference category ($p < 0.05$)

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2015

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version **is also available.**

Cette publication est aussi disponible en français.

Assessment of the quality of the childhood physical abuse measure in the National Population Health Survey

by Margot Shields, Wendy Hovdestad and Lil Tonmyr

Abstract

Background: The long-term health consequences of childhood physical abuse are often studied using retrospective self-reports collected from adults. This study assesses the quality of a question on childhood physical abuse in the National Population Health Survey (NPHS).

Data and methods: All NPHS respondents aged 18 or older (n = 15,027) were asked a question about childhood physical abuse in cycles 1 (1994/1995), 7 (2006/2007) and 8 (2008/2009). The reliability of this question was assessed over these periods. Associations between response patterns to the abuse item and health conditions that are related to childhood physical abuse were examined.

Results: Across all NPHS cycles, very few respondents refused to answer or replied "don't know" to the item on childhood physical abuse. Reliability, as measured by Cohen's kappa statistic, was "substantial" for the two-year interval between cycles 7 and 8, and "moderate" for the 12- and 14-year intervals from cycle 1. Kappa estimates were similar when examined by various demographic factors. Compared with consistent deniers, respondents who consistently affirmed childhood physical abuse and those who provided inconsistent responses had increased odds of depression, fair or poor self-perceived health, disability, migraine, and heart disease.

Interpretation: Despite some limitations, the NPHS question on childhood physical abuse allows researchers to investigate long-term health consequences of abuse.

Keywords: Child abuse, child maltreatment, reliability, retrospective recall, validity

Numerous studies have documented associations between childhood physical abuse and subsequent mental and physical disorders.[1-3] The measure of abuse in these analyses is typically based on retrospective reports from adults. In order to best interpret studies that link childhood physical abuse to health problems in later life, it is important to understand the validity and reliability of such measures.

Validity—the degree to which retrospective responses reflect events that meet objective criteria for physical abuse—is particularly difficult to assess. Poor memory of childhood maltreatment has been documented[4,5] and may result in inaccurate reporting. Because of its sensitive nature, respondents may not report that they experienced abuse. As well, items that are behaviourally specific have higher validity than those that are self-defined or interpretative.[6-9] A review of the literature on the validity of retrospective reporting of adverse childhood experiences indicates that the rate of false negatives is substantial, and that false positives are rare.[7] Research is needed to understand biases in reporting.

Reliability (or consensus) of reports of abuse over time is more easily assessed. As expected, studies requiring shorter-term recall have higher reliability than those with a longer follow-up period.[10-15] And like validity, questions that are behaviourally specific have higher reliability.[7]

Extreme concerns about potential participant upset and consequent item non-response have been noted.[16] If item non-response is an actual problem and if non-respondents and respondents differ in their experiences of abuse, survey results could potentially be biased.

A retrospective question about childhood physical abuse was included in the National Population Health Survey (NPHS), and was asked in cycles 1 (1994/1995), 7 (2006/2007), and 8 (2008/2009). The purposes of this study are to:

1. report item non-response to the NPHS childhood physical abuse question;
2. assess the reliability of the physical abuse item over three periods: 1994/1995 to 2006/2007, 1994/1995 to 2008/2009, and 2006/2007 to 2008/2009;
3. investigate the extent to which demographic factors are related to the reliability of reports of childhood physical abuse;
4. compare the reliability of the childhood physical abuse question with other childhood stressor questions in the NPHS; and
5. examine associations between response patterns to the childhood physical abuse item and depression, fair or poor self-perceived health, disability, migraine and heart disease; that is, to answer the question, "Were consistent affirmers and inconsistent responders to the childhood physical abuse item more likely than consistent deniers to report these conditions?"

Methods

Data source

The analysis is based on longitudinal data from nine cycles (1994/1995 through 2010/2011) of the NPHS. The target population of the NPHS was household residents in the 10 provinces

Authors: Margot Shields, Wendy Hovdestad (wendy.hovdestad@phac-aspc.gc.ca) and Lil Tonmyr (Lil.Tonmyr@phac-aspc.gc.ca) are with the Public Health Agency of Canada.

in 1994/1995, excluding residents of Indian Reserves, institutions, Canadian Forces bases, and some remote areas.

In 1994/1995, 20,095 households were selected for the longitudinal panel. One person in each household was selected at random; of these, 86% (17,276) completed the General component of the questionnaire in 1994/1995. Every two years since then, attempts were made to re-interview these individuals. Data collection ended in 2010/2011, providing 16 years of follow-up. The NPHS "square" file contains one record for each of the 17,276 respondents with data on a comprehensive set of health-related questions asked in each cycle. In cycle 1, 75% of interviews were conducted in person. In later cycles, almost all interviews were by telephone (more than 99% in cycles 7 and 8). Detailed descriptions of the NPHS design, sample, and interview procedures are available in published reports.[17-19]

NPHS respondents were asked for permission to share the information they provided with provincial ministries of health, Health Canada, and the Public Health Agency of Canada. Most respondents (15,631; 90%) agreed to share. This article is based on data from the "square share" file.

The questions on childhood stressors (including physical abuse) were asked in cycles 1 (1994/1995), 7 (2006/2007), and 8 (2008/2009) of all respondents aged 18 or older. The questions were also asked of respondents who were younger than 18 in cycle 1 who had turned 18 by cycle 4 (1,034). For the present analyses, cycle 1 responses were updated to reflect the cycle 4 responses for these cases. This article is based on 15,027 records for which the questions on childhood stressors were applicable in at least one cycle. Because there was no restriction on age when the original sample was selected in 1994/1995, it was possible to refresh the study sample for cycles 7 and 8 to include respondents who had turned 18 by this time. However, people who immigrated to Canada after 1994/1995 are excluded from this study.

For much of the NPHS, proxy responses (usually from a family member) were permitted if the selected respondent was incapacitated by a health problem. Because of their personal nature, this did not apply to the childhood stressor questions. Responses were set to "not stated" for respondents whose information was provided by a proxy reporter.

Measures

Childhood stressors

The question on *childhood physical abuse* was one of seven items in the childhood stressor module. "The next few questions ask about some things that may have happened to you while you were a child or a teenager, before you moved out of the house. Were you ever physically abused by someone close to you?" Table 4 contains the wording of the other six items.

Health conditions

Responses to the childhood physical abuse question were examined in relation to five health conditions shown in the literature to be related to childhood physical abuse.[1-3] Questions about these five health conditions were asked in all nine NPHS cycles; respondents were categorized as having the condition if it was reported in one or more cycles.

Depression was measured using a subset of questions from the Composite International Diagnostic Interview, according to the method of Kessler *et al.*[20]

Self-perceived health was assessed with the question, "In general, would you say your health is: excellent? very good? good? fair? poor?"

Disability was defined based on an affirmative response to any of the following questions: "Because of a long-term physical or mental condition or a health problem, are you limited in the kind or amount of activity you can do: at home? at school? at work? in other activities?" "Do you have any long-term disabilities or handicaps?"

The presence of *chronic conditions* was established by asking respondents if a health professional had diagnosed them as having condition(s) that had lasted, or were expected to last, at least six months. Two chronic conditions were considered in this study: *migraine* and *heart disease*.

Analysis

To examine the reliability of the childhood stressor questions across NPHS cycles, two measures of consensus were calculated: concordance rates (percentage agreement) and Cohen's kappa statistic. Concordance rates are easy to interpret, but a disadvantage is the possibility of artificially inflated agreement rates for variables with a low prevalence because most values fall into one category.[21]

Cohen's kappa corrects for the percentage of agreement that could be expected to occur by chance and is useful when concordance rates may be artificially high due to low prevalence.[21] The major disadvantage of the kappa statistic is that it is difficult to interpret. Landis and Koch[22] proposed the following interpretation, which is commonly used in reliability studies: 0.01-0.20 *slight agreement*; 0.21-0.40 *fair agreement*; 0.41-0.60 *moderate agreement*; 0.61-0.80 *substantial agreement*; and 0.81-0.99 *almost perfect agreement*.

To investigate the extent to which demographic factors are related to the reliability of retrospective reports of childhood physical abuse, estimates of kappa were compared by sex, age group (younger than 25, 25 to 39, 40 to 64, 65 or older), marital status (married/living common-law; yes/no), household income (lowest household income quintile; yes/no), and race (white/non-white). As well, these factors were examined in a logistic regression model with concordance as the dependent variable. A limitation of this approach is that if prevalence is low in a demographic group (for example, married), concordance may be artificially high, and therefore, associations with concordance may be due to low prevalence rather than consistent responses. Consequently, an approach recommended by McKinney[23] was used, where demographic comparisons were made only for respondents with at least one report of abuse (that is, demographic factors were examined for those who

consistently affirmed abuse versus those who gave inconsistent responses).

Logistic regression was also used to examine associations between response patterns to the childhood physical abuse item and health conditions.

Analyses were conducted using SAS Enterprise Guide 5.1. All estimates are based on weighted data. Weights were created at Statistics Canada so that the data would be representative of the Canadian population living in the 10 provinces in 1994/1995, and were inflated to compensate for non-response to the NPHS in 1994/1995. Variance estimates and 95% confidence intervals (CIs) were calculated using the bootstrap technique (with the SAS "proc survey" procedures) to account for the complex survey design.[24] Because SAS does not provide a procedure for estimating the variance of the kappa statistic using the bootstrap technique, CIs for kappa were estimated using a design effect of two to account for the NPHS sample design.

Results

Table 1 shows responses to the childhood physical abuse question across NPHS cycles. The main reasons for non-response were total non-response to the cycle, proxy reporting, and partial non-response to the cycle (interview terminated before childhood stressor questions were asked). Across all cycles, very few respondents refused to answer the childhood physical abuse question or answered "don't know." Percentages responding "don't know" in cycles 1 and 8 and refusing to respond in cycle 7 were too low (fewer than 20 cases) to report based on NPHS release guidelines.

About 8% of respondents reported childhood physical abuse in each NPHS cycle. Based in a report of abuse in any cycle, the prevalence rose to 11%. Women were more likely than men to report physical abuse; reporting physical abuse was less likely among older respondents (Appendix Table A).

The majority of respondents answered the childhood physical abuse item in two or more cycles: 40% provided a response in three cycles; 18%, in two cycles; 35% in one cycle; and 8% had no responses to the question (data not shown). In cases where answers were inconsistent, transitions from "no" to "yes" were slightly more common than transitions from "yes" to "no" (Table 2), but only between cycles 1 and 8 was the difference between these two transitions (4.0% versus 2.9%) statistically significant.

Concordance rates and kappa statistics were calculated for respondents who provided two or more responses to the physical abuse item (Table 3). Based on the guidelines of Landis and Koch,[22] the kappa statistic indicates a "moderate" level of agreement in responses between cycles 1 and 7 (12-year interval) and cycles 1 and 8 (14-year interval). Between cycles 7 and 8 (2-year interval), agreement is "substantial." Likewise, the percentage of respondents providing consistent answers was slightly higher for the 2-year interval between cycles 7 and 8 than for the longer periods involving comparisons with cycle 1. Among respondents who provided answers in all

Table 1
Response status and weighted percentages for National Population Health Survey childhood physical abuse question, by survey cycle

	Cycle 1†				Cycle 7				Cycle 8			
	1994/1995 or 2000/2001				2006/2007				2008/2009			
	n	%	%‡	%§	n	%	%‡	%§	n	%	%‡	%§
Total	15,027				15,027				15,027			
Total responding to childhood physical abuse question	12,786	79.8	90.3	100.0	8,752	57.0	65.5	100.0	8,165	53.5	60.5	100.0
Yes	1,056	6.2	7.0	7.8	771	4.8	5.5	8.4	704	4.3	4.9	8.1
No	11,730	73.6	83.3	92.2	7,981	52.2	60.0	91.6	7,461	49.2	55.6	91.9
Non-response												
Total	1,038	8.6	9.7	...	3,993	30.0	34.5	...	4,559	34.9	39.5	...
Full non-response to cycle	158	1.6	1.8	...	3,109	24.5	28.1	...	4,062	31.7	35.8	...
Partial non-response to cycle	359	2.9	3.2	...	307	1.9	2.1	...	97	0.6	0.7	...
Non-response due to proxy reporting	490	3.9	4.4	...	349	2.4	2.8	...	290	2.0	2.3	...
Item non-response - don't know	F	F	F	...	213	1.2	1.3	...	F	F	F	...
Item non-response - refusal††	21	0.2E	0.2E	...	F	F	F	...	89	0.5	0.6	...
Not applicable												
Total	1,203	11.6	...	...	2,282	13.0	...	...	2,303	11.5	...	...
Younger than 18	1,203	11.6	...	...	272	3.0	...	...	0	...	...	...
Institution	0	...	...	...	119	0.5	...	...	106	0.5	...	...
Dead	0	...	...	...	1,891	9.5	...	...	2,197	11.1	...	...

† respondents younger than 18 in cycle 1 who had turned 18 by cycle 4 (n = 1,034) were asked childhood physical abuse question in cycle 4
‡ excluding not applicable
§ excluding not applicable and non-response
†† respondents who refused to answer first question in childhood stressor module skipped out of module; refusal counts include these cases and respondents who continued but refused to answer childhood physical abuse question
... not applicable
E use with caution
F too unreliable to be published
Source: 1994/1995 to 2010/2011 National Population Health Survey (square, share file).

Table 2
Response patterns across National Population Health Survey cycles to childhood physical abuse question

Survey cycles/Response pattern across survey cycles	n	Weighted percentage	95% confidence interval	
		%	from	to
Cycle 1 - Cycle 7 (n = 7,616)				
Yes/Yes	396	5.0	4.2	5.8
Yes/No	266	2.9	2.5	3.4
No/Yes	294	3.6	3.1	4.1
No/No	6,660	88.4	87.4	89.4
Cycle 1 - Cycle 8 (n = 6,980)				
Yes/Yes	339	4.5	3.9	5.1
Yes/No	246	2.9	2.5	3.4
No/Yes	283	4.0	3.4	4.6
No/No	6,112	88.6	87.6	89.6
Cycle 7 - Cycle 8 (n = 7,385)				
Yes/Yes	446	5.7	5.1	6.4
Yes/No	183	2.2	1.8	2.6
No/Yes	196	2.5	2.0	2.9
No/No	6,560	89.6	88.7	90.5

Source: 1994/1995 to 2010/2011 National Population Health Survey (square, share file).

Table 3
Agreement across National Population Health Survey cycles to childhood physical abuse question

Survey cycles	Kappa			Concordance		
	K	95% confidence interval		Rate	95% confidence interval	
		from	to	%	from	to
Cycle 1 - Cycle 7 (n = 7,616)	0.57	0.52	0.62	93.4	92.8	94.1
Cycle 1 - Cycle 8 (n = 6,980)	0.53	0.47	0.58	93.1	92.4	93.9
Cycle 7 - Cycle 8 (n = 7,385)	0.69	0.64	0.73	95.3	94.8	95.9
Cycle 1 - Cycle 7 - Cycle 8 (n = 6,501)	...	...	...	91.0	90.1	91.8

... not applicable
Source: 1994/1995 to 2010/2011 National Population Health Survey (square, share file).

three cycles, more than 90% were consistent across all three.

Comparisons of reliability estimates were made by sex, age group, marital status, household income, and race (data not shown). Based on kappa, reliability estimates were similar for all covariates. Similarly, among respondents with at least one report of childhood physical abuse, logistic regression results indicated that the likelihood of consistently affirming abuse was not related to demographic factors (data not shown).

Table 4 shows reliability estimates for the other items in the childhood stressor module. Based on kappa, reliability estimates were higher for the items on parental divorce and parental alcohol/drug problems, and lower for the items on father/mother not having a job when they wanted to be working, a frightening experience, and being sent away for doing something wrong.

Associations between response patterns to the physical abuse item and health conditions that have been shown to be related to childhood physical abuse are presented in Table 5. In all cases, respondents who provided inconsistent responses had higher odds of reporting health conditions than those who consistently denied that physical abuse occurred. Compared with people who consistently responded "yes," those who said "no" in two cycles and "yes" in one cycle had somewhat lower odds of reporting health conditions.

Discussion

This study, based on three cycles of data collected between 1994/1995 and 2008/2009 by the National Population Health Survey, indicates that 91% of respondents provided consistent answers about childhood physical abuse across all cycles; very few refused to answer or replied "don't know."

Reliability, as measured by Cohen's kappa statistic, was "substantial" for the two-year interval between cycles 7 and 8 and "moderate" for the 12- and 14-year intervals from cycle 1. Compared with consistent deniers, consistent affirmers and inconsistent responders to the childhood physical abuse item had increased odds of reporting depression, fair or poor self-perceived health, disability, migraine, and heart disease.

The percentages reporting that they had experienced childhood physical abuse was similar in each cycle—8%—with 11% reporting physical abuse in at least one cycle. A 2012 cross-sectional survey of the Canadian population[1] yielded a substantially higher estimate of retrospectively reported childhood physical abuse (26%), which was similar to estimates from a 1990 Ontario survey.[25] The phrase "someone close to you" in the NPHS question may have resulted in the exclusion of incidents where the perpetrator was a teacher or other individual to whom the respondent did not feel "close." It would also exclude physical assaults committed by strangers. As well, unlike the NPHS, the other surveys asked a series of questions on specific events. NPHS estimates are more in line with estimates from the 2012 and 1990 surveys based on questions that used terms indicative of severe abuse (for example, "kick," "punch," "choke" or "burn"). A study that used data from two large, community-based surveys conducted in the United States in 1997 and 2003 found that 8% of respondents labelled themselves as being "physically abused," while 16% reported at least one behaviourally defined physical abuse event.[9]

Generally, retrospective questions that focus on behaviourally specific events

Table 4
Agreement across National Population Health Survey cycles to childhood stressor questions

Childhood stressor questions	Percentage "yes"			Cycle 1 - Cycle 7 Kappa			Cycle 1 - Cycle 8 Kappa			Cycle 7 - Cycle 8 Kappa		
					95% confidence interval			95% confidence interval			95% confidence interval	
	Cycle 1	Cycle 7	Cycle 8	K	from	to	K	from	to	K	from	to
Were you ever physically abused by someone close to you?	7.8	8.4	8.1	0.57	0.52	0.62	0.53	0.47	0.58	0.69	0.64	0.73
Did you spend two weeks or more in the hospital?	15.5	14.8	13.9	0.52	0.49	0.56	0.52	0.48	0.56	0.63	0.59	0.67
Did your parents get a divorce?	12.5	14.7	15.0	0.80	0.77	0.83	0.79	0.76	0.82	0.90	0.88	0.92
Did your father or mother not have a job for a long time when they wanted to be working?	13.8	11.2	12.1	0.39	0.34	0.43	0.39	0.34	0.43	0.52	0.47	0.57
Did something happen that scared you so much that you thought about it for years after?	21.8	19.5	19.6	0.39	0.35	0.43	0.38	0.34	0.42	0.55	0.51	0.58
Were you sent away from home because you did something wrong?	2.6	2.0	2.1	0.41	0.30	0.52	0.44	0.33	0.55	0.54	0.43	0.64
Did either of your parents drink or use drugs so often that it caused problems for the family?	14.4	14.1	14.6	0.66	0.63	0.70	0.63	0.60	0.67	0.77	0.74	0.80

Source: 1994/1995 to 2010/2011 National Population Health Survey (square, share file).

Table 5
Adjusted odds ratios relating response patterns to National Population Health Survey childhood physical abuse question to selected health-related conditions

Health-related condition/Response pattern to childhood physical abuse question across survey cycles	Adjusted[†] odds ratio	95% confidence interval	
		from	to
Depression			
All yes (three yes, or two yes and one missing)	4.4*	3.2	6.0
Two yes and one no	5.0*	3.6	6.9
Two no and one yes	3.2*	2.5	4.2
One yes and one no (and one missing)	2.2*	1.4	3.3
All no (three no or two no and one missing)[‡]	1.0	...	...
Fair or poor self-perceived health			
All yes (three yes or two yes and one missing)	2.5*	1.8	3.4
Two yes and one no	2.7*	2.0	3.8
Two no and one yes	1.6*	1.2	2.2
One yes and one no (and one missing)	2.7*	1.7	4.1
All no (three no or two no and one missing)[‡]	1.0	...	...
Disability			
All yes (three yes or two yes and one missing)	3.6*	2.5	5.1
Two yes and one no	2.5*	1.7	3.6
Two no and one yes	1.8*	1.4	2.4
One yes and one no (and one missing)	2.8*	1.8	4.3
All no (three no or two no and one missing)[‡]	1.0	...	...
Migraine			
All yes (three yes or two yes and one missing)	2.9*	2.0	4.1
Two yes and one no	2.1*	1.5	3.0
Two no and one yes	1.8*	1.3	2.4
One yes and one no (and one missing)	1.8*	1.2	2.7
All no (three no or two no and one missing)[‡]	1.0	...	...
Heart disease			
All yes (three yes or two yes and one missing)	1.8*	1.2	2.6
Two yes and one no	1.7*	1.1	2.7
Two no and one yes	1.5*	1.0	2.4
One yes and one no (and one missing)[§]	1.8	1.0	3.4
All no (three no or two no and one missing)[‡]	1.0	...	...

† adjusted for sex, age, marital status, household income, and race
‡ reference group
§ p-value approaches significance ($p = 0.06$)
* significantly different from reference group ($p < 0.05$)
... not applicable
Source: 1994/1995 to 2010/2011 National Population Health Survey (square, share file).

using multiple items yield higher prevalence estimates with higher validity[6-9,26,27] than does the broad NPHS question, which required respondents to subjectively define physical abuse and then decide if past experiences constituted abuse. For example, the NPHS question on parental divorce (a specific event) had higher reliability, while more subjective items such as, "Did your father or mother not have a job for a long time when they wanted to be working?" yielded lower kappa estimates.

In the NPHS, women were more likely than men to report childhood physical abuse. This is contrary to findings from other Canadian studies,[1,25,28] although in two of these studies, estimates for severe physical abuse were similar between the sexes.[25,28] A global meta-analysis involving close to 10 million participants concluded that there were no gender differences in the prevalence of physical abuse.[29] In the NPHS, the self-defined nature of the item and the specification of "someone close to you" may have resulted in the relatively lower estimate among men.

The relationship between childhood physical abuse and premature mortality[30,31] may have contributed to the lower estimates for older respondents.

Similar to the NPHS results, other retrospective studies using a variety of items or groups of items to measure childhood physical abuse have yielded kappa

values in the "moderate" range.[23,32-34] Higher kappa values (at or approaching "substantial") have been observed for studies based on clinical samples[6] and for those with relatively short intervals (less than two years) between measurements,[10,11,13,14] consistent with the higher kappa value (0.69) for the two-year interval between cycles 7 and 8 of the NPHS.

Earlier studies have also found that the reliability of childhood physical abuse estimates was not related to demographic variables.[11,13] However, contrary to NPHS results, McKinney *et al.*[23] reported that among respondents with at least one report of abuse, low education and younger age were associated with inconsistent responses.

Although assessing the *reliability* of the NPHS question on childhood physical abuse is important, the *validity* (or the extent to which responses reflect events that meet objective criteria for physical abuse) is equally or more important. When responses are inconsistent, how should respondents be classified?

The validity of retrospective reports of childhood experiences has been questioned, particularly for people with mood or anxiety disorders.[35-38] However, reviews of the literature support the accuracy of early memories, especially for events such as childhood maltreatment.[7,10] Little or no evidence supports a relationship between anxiety and depression and short- or long-term memory deficits or distortions.[7,10] Furthermore, validity studies have found that false positive reports of abuse are rare, whereas false negatives are fairly common. In well-documented cases of adults who experienced serious childhood abuse, about one-third did not report it.[7]

In light of these studies, it would seem well-advised to classify NPHS respondents with inconsistent reports as having experienced childhood physical abuse. The elevated odds ratios for health conditions known to be associated with physical abuse among those who provided inconsistent responses versus those who consistently responded "no" supports this classification. Other researchers have noted that the estimated prevalence of childhood physical abuse increases considerably when subjects are asked about it more than once; they suggest that negative responses among inconsistent reporters may be less accurate than positive responses.[12,23] Given that false positives are rare and that inconsistent responses are relatively common, in longitudinal surveys such as the NPHS it may be beneficial to re-ask childhood maltreatment questions to increase the accuracy of estimates.

Compared with those who consistently responded "yes" to the abuse question, those who said "no" in two cycles and "yes" in one cycle had somewhat lower odds of reporting health conditions. It has been suggested that individuals who experienced childhood abuse, but who function well in adulthood, may be more likely to forget or fail to report maltreatment,[7] which may explain the somewhat more favourable odds for these individuals (although the odds are higher than for those who consistently reported "no"). Severity of abuse may also play a role. Those who experienced more severe abuse may be more likely to consistently recall and report it, and associations between maltreatment and health conditions tend to be stronger with more severe abuse.[1] Failure to treat inconsistent responses as "yes" may alter associations between physical abuse and health outcomes if the inconsistent responses represent the occurrence of less severe abuse.

Mode of data collection can also influence the validity of childhood physical abuse estimates. Some studies suggest that in-person interviews heighten rapport and trust with respondents, thereby increasing the likelihood of disclosure.[39-41] Other studies have found that estimates based on self-administered questionnaires tend to be higher, which suggests that the anonymity and privacy associated with this mode of inquiry may facilitate disclosure.[26,42,43] The majority of NPHS interviews in cycle 1 were conducted in person, whereas nearly all cycle 7 and 8 interviews were conducted by telephone. The similarity of estimates for abuse across cycles implies that disclosure was not influenced by whether interviews were conducted in person or by telephone. However, the extent to which estimates would be higher if self-administered questionnaires had been used is unknown.

What is already known on this subject?

- Retrospective reports of childhood physical abuse involve considerable measurement error and a substantial rate of false negatives.
- False positive reports of childhood physical abuse are rare.
- Retrospective questions that focus on behaviour-specific abuse have higher validity and reliability than broad, subjectively defined questions.

What does this study add?

- Based on data from the National Population Health Survey, the reliability of the question on childhood physical abuse was in the "substantial" range when examined over a two-year interval, and in the "moderate" range for intervals of 12 and 14 years.
- Inconsistent reports of childhood physical abuse were not related to demographic variables.
- Compared with consistent deniers, consistent affirmers and inconsistent responders to the childhood physical abuse item had increased odds of reporting depression, fair or poor self-perceived health, disability, migraine, and heart disease.

Limitations

Although the validity of the NPHS childhood physical abuse question was partially examined by comparing consistency/inconsistency in responses to questions about health conditions, a more thorough assessment would involve comparisons with a gold standard—that is, independent corroboration with external information known to represent whether

or not respondents' experiences meet recognized criteria for childhood physical abuse. The findings of this study suggest that NPHS respondents with inconsistent reports should be classified as having experienced childhood physical abuse, but the extent to which this would improve the accuracy of estimates is unknown. While false positives are rare,[7] they do occur. As well, it is likely that some respondents who consistently denied that abuse took place, may have experienced abuse. Another limitation of the study is non-response— to the NPHS and to the childhood physical abuse question (including non-response due to proxy reporting).

Conclusion

A major advantage of this study is that it is based on a large sample representative of the general population, whereas most reliability and validity studies have used small clinical or convenience samples. The comprehensive set of demographic and health-related questions in the NPHS made it possible to examine reliability across an array of covariates and to examine consistency/inconsistency of responses to childhood physical abuse in relation to several health conditions that have been shown to be related to childhood maltreatment.

Although the NPHS question on childhood physical abuse has limitations, this study suggests that it is a reliable measure over time and that inconsistent responses likely indicate that abuse did occur. The 16-year follow-up period offers the potential to investigate associations between childhood physical abuse and mortality and morbidity over time. This information can, in turn, assist in the development of child maltreatment and chronic disease prevention activities. ■

Acknowledgements

The assistance of Dr. Patrick Boily of the Centre for Quantitative Analysis and Decision Support is gratefully acknowledged, as is the assistance of Caroline Wallace with the literature review.

References

1. Afifi TO, MacMillan HL, Boyle M, et al. Child abuse and mental disorders in Canada. *Canadian Medical Association Journal* 2014; 186(9): E324-32.
2. Norman RE, Byambaa M, De R, et al. The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. *PLoS Medicine* 2012; 9(11): e1001349.
3. Tonmyr L, Jamieson E, Mery LS, MacMillan HL. The relationship between childhood adverse experiences and disability due to physical health problems in a community sample of women. *Women & Health* 2005; 41(4): 23-35.
4. Scott KM, Smith DR, Ellis PM. Prospectively ascertained child maltreatment and its association with DSM-IV mental disorders in young adults. *Archives of General Psychiatry* 2010; 67(7): 712-9.
5. Williams LM. Recall of childhood trauma: a prospective study of women's memories of child sexual abuse. *Journal of Consulting and Clinical Psychology* 1994; 62(6): 1167-76.
6. Fisher HL, Craig TK, Fearon P, et al. Reliability and comparability of psychosis patients' retrospective reports of childhood abuse. *Schizophrenia Bulletin* 2011; 37(3): 546-53.
7. Hardt J, Rutter M. Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *The Journal of Child Psychiatry* 2004; 45(2): 260-73.
8. Monteiro IS, Maia A. Family childhood experiences reports in depressed patients: comparison between 2 time points. *Procedia Social and Behavioral Sciences* 2010; 5: 541-7.
9. Thombs BD, Bernstein DP, Ziegelstein RC, et al. An evaluation of screening questions for childhood abuse in 2 community samples: implications for clinical practice. *Archives of Internal Medicine* 2006; 166(18): 2020-6.
10. Bernstein DP, Fink L, Handelsman L, et al. Initial reliability and validity of a new retrospective measure of child abuse and neglect. *The American Journal of Psychiatry* 1994; 151(8): 1132-6.
11. Dube SR, Williamson DF, Thompson T, et al. Assessing the reliability of retrospective reports of adverse childhood experiences among adult HMO members attending a primary care clinic. *Child Abuse & Neglect* 2004; 28(7): 729-37.
12. Fergusson DM, Horwood LJ, Woodward LJ. The stability of child abuse reports: a longitudinal study of the reporting behaviour of young adults. *Psychological Medicine* 2000; 30(3): 529-44.
13. Friedrich WN, Talley NJ, Panser L, et al. Concordance on reports of childhood abuse by adults. *Child Maltreatment* 1997; 2(2): 164-71.
14. Kooiman CG, Ouwehand AW, ter Kuile MM. The Sexual and Physical Abuse Questionnaire (SPAQ). A screening instrument for adults to assess past and current experiences of abuse. *Child Abuse & Neglect* 2002; 26(9): 939-53.
15. Meyer IH, Muenzenmaier K, Cancienne J, Struening E. Reliability and validity of a measure of sexual and physical abuse histories among women with serious mental illness. *Child Abuse & Neglect* 1996; 20(3): 213-9.
16. Hovdestad W, Tonmyr L. Proposal of childhood maltreatment questions for inclusion on a Canadian national health survey: Notes from the process. In: Gray J, editors. *ISPCAN World Perspectives on Child Abuse*. Aurora, Colorado: International Society for Prevention of Child Abuse and Neglect, 2010: 43-8.
17. Statistics Canada. National Population Health Survey, Household Component, Cycles 1 to 9 (1994/1995 to 2010/2011), *Longitudinal Documentation*. Ottawa: Statistics Canada, 2012.
18. Swain L, Catlin G, Beaudet MP. The National Population Health Survey--its longitudinal nature. *Health Reports* 1999; 10(4): 69-82.
19. Tambay J-L, Catlin G. Sample design of the National Population Health Survey. *Health Reports* 1995; 7(1): 29-38.
20. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry* 1994; 51(1): 8-19.
21. Stemler SE. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation* 2004; 9(4): 1-19.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1): 159-74.

23. McKinney CM, Harris TR, Caetano R. Reliability of self-reported childhood physical abuse by adults and factors predictive of inconsistent reporting. *Violence and Victims* 2009; 24(5): 653-68.

24. Yeo D, Mantel H, Liu TP. *Bootstrap Variance Estimation for the National Population Health Survey*. Proceedings of the Annual Meeting of the American Statistical Association, Survey Research Methods Section, August 1999. Baltimore, Maryland: American Statistical Association, 1999.

25. MacMillan HL, Fleming JE, Trocme N, et al. Prevalence of child physical and sexual abuse in the community. Results from the Ontario Health Supplement. *Journal of the American Medical Association* 1997; 278(2): 131-5.

26. Ji K, Finkelhor D, Dunne M. Child sexual abuse in China: a meta-analysis of 27 studies. *Child Abuse & Neglect* 2013; 37(9): 613-22.

27. Stoltenborgh M, van Ijzendoorn MH, Euser EM, Bakermans-Kranenburg MJ. A global perspective on child sexual abuse: meta-analysis of prevalence around the world. *Child Maltreatment* 2011; 16(2): 79-101.

28. MacMillan HL, Tanaka M, Duku E, et al. Child physical and sexual abuse in a community sample of young adults: results from the Ontario Child Health Study. *Child Abuse & Neglect* 2013; 37(1): 14-21.

29. Stoltenborgh M, Bakermans-Kranenburg MJ, van Ijzendoorn MH, Alink LR. Cultural-geographical differences in the occurrence of child physical abuse? A meta-analysis of global prevalence. *International Journal of Psychology* 2013; 48(2): 81-94.

30. Brown DW, Anda RF, Tiemeier H, et al. Adverse childhood experiences and the risk of premature mortality. *American Journal of Preventive Medicine* 2009; 37(5): 389-96.

31. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine* 1998; 14(4): 245-58.

32. Fergusson DM, Horwood LJ, Boden JM. Structural equation modeling of repeated retrospective reports of childhood maltreatment. *International Journal of Methods in Psychiatric Research* 2011; 20(2): 93-104.

33. Nelson EC, Lynskey MT, Heath AC, et al. A family study of adult twins with and without a history of childhood abuse: stability of retrospective reports of maltreatment and associated family measures. *Twin Research and Human Genetics* 2010; 13(2): 121-30.

34. Yancura LA, Aldwin CM. Stability and change in retrospective reports of childhood experiences over a 5-year period: findings from the Davis Longitudinal Study. *Psychology and Aging* 2009; 24(3): 715-21.

35. Gerlsma C, Emmelkamp PMG, Arrindell WA. Anxiety, depression, and perception of early parenting: a meta-analysis. *Clinical Psychology Review* 1990; 10(3): 251-77.

36. Kendall-Tackett K, Becker-Blease K. The importance of retrospective findings in child maltreatment research. *Child Abuse & Neglect* 2004; 28(7): 723-7.

37. Raphael KG, Widom CS, Lange G. Childhood victimization and pain in adulthood: a prospective investigation. *Pain* 2001; 92(1-2): 283-93.

38. Widom CS, Weiler BL, Cottler LB. Childhood victimization and drug abuse: a comparison of prospective and retrospective findings. *Journal of Consulting and Clinical Psychology* 1999; 67(6): 867-80.

39. da Silva SS, da Costa MA. The stability of self-reported adverse experiences in childhood: a longitudinal study on obesity. *Journal of Interpersonal Violence* 2013; 28(10): 1989-2004.

40. Martin J, Anderson J, Romans S, et al. Asking about child sexual abuse: methodological implications of a two stage survey. *Child Abuse & Neglect* 1993; 17(3): 383-92.

41. Peters SD, Wyatt GE, Finkelhor D. Prevalence. In: Finkelhor D, editors. *A Sourcebook on Child Sexual Abuse*. Beverly Hills, California: Sage, 1986: 15-59.

42. Bebbington PE, Jonas S, Brugha T, et al. Child sexual abuse reported by an English national sample: characteristics and demography. *Social Psychiatry and Psychiatric Epidemiology* 2011; 46(3): 255-62.

43. Dill DL, Chu JA, Grob MC, Eisen SV. The reliability of abuse history reports: a comparison of two inquiry formats. *Comprehensive Psychiatry* 1991; 32(2): 166-9.

Appendix

Table A

Pecentage reporting childhood physical abuse in at least one National Population Health Survey cycle, by sex and age group, household population aged 18 or older in 1994/1995, Canada excluding territories

		95% confidence interval	
	%	from	to
Total	10.6	9.9	11.3
Sex			
Men†	6.8	6.0	7.6
Women	14.2*	13.1	15.3
Age group (years)			
18 to 24	11.1	9.6	12.6
25 to 34	13.5	12.0	15.0
35 to 49†	12.7	11.4	13.9
50 to 64	8.2*	6.9	9.6
65 or older	3.4*	2.5	4.3

† reference group
* significantly different from reference group ($p < 0.05$)
Source: 1994/1995 to 2010/2011 National Population Health Survey (square, share file).